



King's Research Portal

DOI:

[10.1007/978-94-010-0019-2_9](https://doi.org/10.1007/978-94-010-0019-2_9)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Ginzburg, J. (2003). Disentangling Public from Non-public Meaning. In *Current and New Directions in Discourse and Dialogue* (pp. 183 - 212). (Text, Speech and Language Technology; Vol. 22). Springer Netherlands. https://doi.org/10.1007/978-94-010-0019-2_9

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Chapter 1

DISENTANGLING PUBLIC FROM NON-PUBLIC MEANING

Jonathan Ginzburg*

Dept of Computer Science

King's College, London

ginzburg@dcs.kcl.ac.uk

Abstract

Analyses of interaction need to characterize not solely ‘success conditions’, a traditional and important means of analyzing action, but also ‘clarification potential’, the range of potential clarification requests (CRs) available in the aftermath of a conversational move. After briefly considering the very productive and effective ways of producing CRs relating to the grammatically governed content of an utterance, I turn to CRs that pertain to a conversational participant’s non-public intentions, the commonest being the bare *Why?*, dubbed here *Why_{meta}*. I demonstrate that *Why_{meta}* shows distinct behaviour from CRs that pertain to grammatically governed content. The most prominent feature perhaps being that, whereas the latter are almost invariably adjacent to the utterances whose clarification they seek, non-adjacency is quite natural for *Why_{meta}*. It can occur at a stage where a second part adjacency pair response has been provided to the utterance it pertains to, suggesting that the information *Why_{meta}* is seeking is a ‘useful extra’, not an essential ingredient required for providing an appropriate response. Rather than treat *Why_{meta}* as clarifying a contextually instantiable goals/plan parameter, I propose that it be treated as an instance of a metadiscursive utterance like *I don’t want to talk about this*.

Keywords:

Dialogue, Clarification Request, Plan Recognition, Grounding

*The research described here is funded by grant number R00022269 from the Economic and Social Research Council of the United Kingdom and by grant number GR/R04942/01 from the Engineering and Physical Sciences Research Council of the United Kingdom. Portions of this paper have been presented at Stuttgart and Göteborg. For many helpful comments I would like to thank Robin Cooper, Jan van Kuppevelt, Staffan Larsson, and David Traum.

1. Introduction

Specifying communicative actions in terms of preconditions and effects is standard. However, when considering communicative *interactions* there is an additional aspect that needs to be considered. Communication is frequently less than perfect. The effect a speaker hopes to achieve doesn't quite go through. And yet in the context of a dialogue, this does not result in *failure*, let alone a crashing halt. Rather, without further ado the (previous) addressee may pose a clarification request (CR), in which she localizes her problem with the utterance. Hence, the aspect additional to intended effect we need analyses of interaction to capture is the class of potential CRs. In this paper I would like to illustrate how characterising certain classes of CRs can shed light on a familiar problem—distinguishing publicly expressed (overt) communicative effects from non-public (implicit) ones.

(1a,b) and (2a,b), all taken from the British National Corpus (BNC),¹ exemplify two quite similar cases of CRs: in each case the initial speaker falsely assumes that an entity (a time in (1a), a location in (1b)) or attribute (as in (2a,b)) they're referring to is obvious ; in each case the addressee poses a CR that pinpoints the referential problem. The pinpointing is effected by using ellipsis, which I will refer to as Clarification Ellipsis (CE), which establishes parallelism with the problematic, to-be-clarified subutterance. Indeed, in context Wendy could express the content she conveys in (2b) even more succinctly as (2c):

- (1) a. Geoffrey: What time did you get up?
Jonathan: When?
Geoffrey: This morning. (BNC)
- b. Dorothy: Oh look, somebody's chopped a daffodil head off!
Christopher: Where?
Dorothy: Here look. (BNC)
- (2) a. Cherrilyn: Yeah I mean (pause) dog hairs rise anyway so
Fiona: What do you mean, rise?
Cherrilyn: The hair (pause) it rises upstairs. (BNC)
- b. Bev: I'm safe.
Wendy: What do you mean I'm safe?
Bev: In case of a fire. (BNC,)

¹All examples from the BNC in this paper were found using Matt Purver's search engine SCORE (Purver, 2001).

c. Wendy: safe?

(Purver et al.,) point out that CRs are almost invariably adjacent to the utterance whose clarification they request.² This generalization derives, presumably, to a large extent from the fact that resolving the referential parameters of an utterance is typically a necessary condition for understanding and appropriately reacting to the utterance, i.e. *grounding* the utterance (in the sense of (Clark, 1996)). Indeed, one could, somewhat simple-mindedly, propose a model of the process which leads to grounding or to the posing of a CR of an utterance based on this idea: the addressee tries to instantiate parameters projected by the various sub-utterances of the utterance. Ability to instantiate leads to grounding. Inability to instantiate one or more parameter triggers the posing of a CR, constructed on the basis of certain operations which take as input the *linguistic sign* associated with the utterance. Such a grounding-as-parameter-instantiation (GPI) model of grounding/CR-triggering has indeed been proposed in (Ginzburg and Cooper, 2001; Ginzburg and Cooper,) and used to provide a linguistic analysis of CRs such as (1), (2).³

This view of grounding is oriented exclusively to what one might call the literal meaning of utterances, those aspects of meaning that derive from grammatically governed processes. One obvious issue that confronts this view, in particular as a theory of CR-triggering, is how to accommodate examples such as (3):

- (3) a. Cherrilyn: You got a pound?
 Jessica: Why?
 Cherrilyn: ch I mean in change
 Jessica: no. (BNC)

b. Why do you ask if I got a pound?

²Note though that UC CRs need not be immediately adjacent, but can be 'stacked', as in the following example:

- (i) Nick: Richard hit the ball on the car.
 Terry: What car?
 Nick: The car that was going past.
 Terry: What ball?
 Nick: James [last or full name] 's football. (BNC)

³The fact that operations specifying the emergence of CRs are defined on signs rather than solely semantic representations is motivated *inter alia* by syntactic and phonological parallelism phenomena.

Jessica's *why?* gets resolved as (3b), a resolution I shall henceforth refer to as *Why_{meta}*.⁴ A prosaic reason for being concerned with *Why_{meta}* is that it is, relatively speaking, high frequency.⁵

In terms of surface syntax *Why_{meta}* looks very much like *When?* and *Where?* from (1). Semantically, however, they do differ in a number of important ways: *when*, *where* and indeed virtually all other *wh*-phrases used in CRs query a constituent of what one might call the nucleus of a given content (bold face in (4b)), whereas *Why_{meta}* outscopes this nucleus to query at the illocutionary level (bold face in (4d)):

- (4) a. A: Did Bo leave? B: Who?/When?/Where?
- b. $\lambda b/t/l Ask(A, ?\mathbf{Leave}(b, t, l))$
- c. A: Did Bo leave? B: Why?
- d. $\lambda r.Cause(r, \mathbf{Ask}(\mathbf{A}, ?\mathbf{Leave}(b, t, l)))$

Be that as it may,⁶ perhaps the most central distinction is that *Why_{meta}* queries for information that is not directly communicated in an utterance, in contrast to the other CRs we saw above. Hence, the challenge posed by *Why_{meta}* to a GPI view of grounding/CR-triggering is how to account for its coherence as a CR given that its antecedent utterance involves no (grammatically provided) contextual parameter. Indeed, an obvious impulse is to add an *intentional* component into the picture, e.g. by somehow projecting a contextual parameter for a plan/goal.⁷ This is a lesson which should be obvious enough given twenty odd years of AI work on dialogue. Postulating goals and plans in order to explain cooperative behaviour was made popular by the Toronto school (including Allen, Cohen, Levesque, Perrault). They showed how this strategy could offer a computational account of a number of phenomena involving what seem to be deviations from 'literal behaviour'. As Allen and Perrault put it already in 1980:

⁴We will see examples of such resolutions where the antecedent utterance is an assertion or command, in which case an appropriate paraphrase is *why do you claim/believe ...*, *Why do you order me to ...*

⁵In the BNC there are 1026 occurrences of interrogatives of the form *Why?*, which constitutes roughly 53% of all *why* interrogatives, 6% of all one word queries and 0.8% of all queries. Of these *Why?*'s, a random sample from one block of the BNC suggests that approximately 30% are *Why_{meta}*. By comparison, the corresponding figures for *Where?* are: 252 occurrences, which is 24% of all *where* interrogatives or 1% of all one word queries; estimate of CR percentage among *Where?* is 60%.

⁶There are a number of additional distinctions between *why* and other modifier *wh*-phrases which lead (Ginzburg, 2001b) to propose that the former is a functor on *facts* as , whereas the others are functors on verbal denotations.

⁷Indeed the GODIS system (Cooper et al., 2000) does something very much like this.

We assume that agents attempt to recognize the plans of other agents and, then, use this plan when deciding what response to make. (from the abstract of (Allen and Perrault, 1980))

One way of reading this is that utterance understanding involves, to a significant extent, recognition of the speaker's goals in making the utterance. So that goal resolution is part and parcel of grounding/CR-triggering. This much by now seems almost self evident—it could be described as a core assumption of computational design of dialogue systems.

Corpus investigation, nonetheless, does occasionally unearth slightly surprising data. As I mentioned above, there is solid evidence for the (none too surprising) generalization that CRs are adjacent to the utterance that serves as their trigger. But what one finds readily enough in a corpus like the BNC is quite different behaviour exhibited by *Why_{meta}*. In the examples of (5), *Why_{meta}* takes as its antecedent an utterance that is for all intents and purposes grounded, that has occurred a number of turns back.

- (5) a. Norrine: When is the barbecue, the twentieth? (pause) Something of June
 Chris: Thirtieth.
 Norrine: A Sunday.
 Chris: Sunday.
 Norrine: Mm.
 Chris: Why? (= *Why do you ask when the barbecue is*)
 Norrine: Becau Because I forgot (pause) That was the day I was thinking of having a proper lunch party but I won't do it if you're going out. (BNC)
- b. Michael: When do you go down town?
 Wendy: Friday.
 Michael: Oh, not before?
 Wendy: Well I don't know. Why? (= *why are you asking when I'm going down town?* or *why are you asking if I won't be going before Friday?*) (BNC)
- c. Cherrilyn: Are you still (pause) erm (pause) going to Bristol (pause) on Monday?
 Fiona: Dunno.
 Cherrilyn: No?
 Fiona: I dunno. Doubt it, why? (= *Why do you ask if I'm going to Bristol?*)
 Cherrilyn: I just wondered.

Fiona: Why?

Cherrilyn: I just wondered! (BNC)

The implication of such examples is this: a CP B can respond coherently to a CP A's utterance **without considering why A made her utterance**. Note also that a CE is quite inappropriate at this stage:

- (6) Norrine: When is the barbecue, the twentieth? (pause) Something of June
 Chris: Thirtieth.
 Norrine: A Sunday.
 Chris: Sunday.
 Norrine: Mm.
 Chris: #(What do you mean) when is the barbecue?

The contrasts exhibited in (6) and (5) are instances of a fundamental bifurcation concerning utterance content and grounding/CR-triggering. Below I will provide some more data that attempts to establish the following:

- A distinction needs to be made between a notion of *utterer's content* (UC), tightly constrained by the (literal) meaning of an utterance, and the goals/plan associated with the utterance by its speaker *utterer's plan* (UP).⁸
- **Recognizing utterer content is part and parcel of the grounding process of an utterance; utterer's plan recognition is not: no presupposition of the goals/plan associated with an utterance emerges after grounding.**

Related to this, the most immediate aim of this paper will be to develop an analysis, based on data from the BNC, of how *Why_{meta}* gets resolved in context and implications this has for the nature of interaction between agents in dialogue. I will compare this resolution process with the resolution process involved in CE, which I will sketch based on (Ginzburg and Cooper,).

The paper is structured as follows: in section 2 I provide some additional data reinforcing the distinction between UC and UP. In section 3,

⁸The term 'utterer's content' is inspired by Grice's term 'speaker meaning', but will be more narrowly circumscribed; also 'content' is used rather than 'meaning' because, following the traditions of California semantics (Montague, Kaplan, Barwise and Perry), we use the latter as a notion at the level of types, whereas the former at the level of tokens.

I provide some background notions from the KOS framework required for my formalization and apply these to explicate UC. In section 4 I consider a previous analysis of *Why_{meta}*, that of (Moore, 1995) and develop my own analysis, which involves viewing it as part of the arsenal of *metadiscursive* utterances available to CPs. Finally in section 5 I draw some conclusions.

2. Utterer's Content v. Utterer's Plan

2.1 Data

I start by considering certain data which indicate the different status the utterer's content has from the utterer's plan.

Antecedents for CR derive from UC, not UP. As exemplified in (3), clarification can target underlying goals of or reasons for utterance. Interestingly, *Why_{meta}* targets only the directly/publicly conveyed content:

- (7) a. [A is a ticket tout, standing outside a theatre] A: It's the best show in London. I have a ticket for tonight's performance. B: (But) Why? (\neq Why do you want to sell your ticket).
- b. A: Do you want some coffee? B: Coffee would keep me awake. A: Why? (\neq Why do you not want coffee?).

UP does not give rise to CE clarification:

- (8) a. A: I've got a ticket for tonight's show. (\rightarrow I want to sell a ticket for tonight's show.) B: You/A ticket/ tonight's show? B: #You want?/
- b. B: Shall I make a coffee? A: Coffee will keep me awake. (\rightarrow I don't want coffee) B: Awake? B: #You don't?

UC recognition is a commitment of grounding. (9) illustrate that by grounding an utterance, presuppositions emerge as to the referential anchoring of contextual parameters. These undermine attempts to pose CRs.⁹

⁹Of course these presuppositions are defeasible. In (i) B suddenly realizes his initial resolution of 'Jo' might be incorrect:

- (i) A: Did Jo leave?
 B: I don't know.
 A: I see.
 B: Wait, which Jo are we talking about?

- (9) a. A: Did Jo leave?
 B: I don't know.
 A: I see.
 B: Anyway, # what do you mean by leave?
 A: Jo left.
 B: I see.
 A: He was displeased with the management.
 B: hmm, # who is Jo?

UP recognition is not a precondition for or a commitment of grounding. In (5) above we saw three examples where CPs (Chris/Wendy/Fiona) each respond to a query and then a number of turns following that pose a *Why_{meta}* about the original query. As I pointed out there, the implication of such examples is this: a CP B can respond coherently to a CP A's utterance **without considering why A made her utterance** (beyond the desire to convey UC). Indeed as the constructed example based on (5a) shows there really is no commitment to B finding out nor A making clear why A made her utterance:

- (10) Norrine: When is the barbecue, the twentieth? (pause) Something of June
 Chris: Thirtieth.
 Norrine: A Sunday.
 Chris: Sunday.
 Norrine: Mm. You know why I'm asking?
 Chris: No. Why?

Dialogue coherence wins over private intentional coherence.

A final point to note w/r to UC v. UP concerns the notion of an incoherent utterance, i.e. one which licenses a CR concerning its basic motivation. An utterance which is dialogically coherent but whose UP is unclear does not allow for such clarification; it only allows for clarification of its UC. Thus, in (11), B provides only a partial answer to A's query. Nonetheless, the infelicity of CRs such as (11c) show that the dialogical coherence of B's response is guaranteed, given that it is at least a partial answer to the question posed by A. The clarification that can be requested is for the UC:¹⁰

- (11) a. A: Where is the train leaving from?

¹⁰Evidence from the BNC suggests that on the whole CRs of the form *What do you mean?* (i.e. without explicit reference to a problematic subutterance) are, nonetheless, construed as clarifications related to the explicitly enunciated content.

- b. B: Somewhere over there.
- c. A: #Why? / #Why do you say this? / #And so?
- d. A: Somewhere over there?/What do you mean somewhere over there?

2.2 Interim Conclusions

In the previous section I presented evidence about the difference evinced by two types of ‘contents’ one can associate with an utterance, the utterer’s content (UC) and the utterer’s plan (UP). The evidence suggests that ‘Why’ does not clarify a parameter akin to CE.

- Recognizing UC is a commitment of grounding; CRs concerning UC are based on a highly localized mechanism whose input is the entire *sign* associated with the utterance.
- Recognizing UP is not a commitment of grounding; CR concerning UP is based on a mechanism which is strictly less localized than the CR mechanism that targets UC.

In the remainder of the paper I need to sketch a picture of how utterances update CP’s information states which accords with this evidence. This picture must, in particular, underpin the contextual background to CRs like (1), (2), (5) above. The essence of the proposed picture is as follows:

- 1 Grounding interaction revolves around the UC, a fleshing out of the utterance skeleton associated with the utterance’s sign.
- 2 There exist dialogue rules for effecting the conversation at the metadiscursive level. These rules regulate *inter alia* utterances such as *Why_{meta}*.

3. Clarifying Utterer’s Content

In this section I sketch an analysis of how CRs of the utterer’s content arise and provide necessary background for the account of *Why_{meta}* in section 4.

3.1 Context in Dialogue

I assume the situation semantics based theory of dialogue context developed in the KOS framework (Ginzburg, 1996; Ginzburg, forthcoming; Bohlin et al., 1999; Larsson, 2002). In ((Ginzburg, 1997b; Ginzburg,

1997a; Ginzburg, 2001b)) I demonstrated the existence of intrinsic asymmetries in context between speaker and addressee w/r to ellipsis resolution of bare wh-phrases. In (12a), ‘why’ must pick up on a fact that positively resolves the initial question A poses, whereas when ‘why’ is uttered by a new speaker, as in (12b), the resolution is *Why_{meta}*. Note that these data cannot be explained merely as a consequence of the differing coherence of an utterance depending on who makes the utterance: the resolution unavailable to A in (12a) *is* coherent and entirely plausible when it arises from a non-elliptical utterance, whose resolution is not so heavily reliant on context, as in (12c):

- (12) (a) A: Where was your Grandmother’s sister born? Why? (Unambiguously: ‘Why was she born *there*?’)
 (b) A: Where was your Grandmother’s sister born? B: Why? (‘Why do you ask where she was born?’)
 (c) A: Where was your Grandmother’s sister born? (and) Why am I asking this question?

Phenomena such as this, in which one CP’s contextual possibilities are distinct from another CP’s, suggest that a single “context” is not fully adequate to describe dialogue, even when talking about “public” context, which results from overtly registered conversational actions. The approach to context common in formal semantics following Stalnaker ((Stalnaker, 1978)) needs to be recast somewhat so that the state of the dialogue at a given point is given in terms of the collection of individual information states of the CPs. This does not necessitate a solipsistic approach, given the considerable evidence, both semantic and psycholinguistic, that CPs try to maintain a common view of the conversation and its background. This, along with other important insights, is captured by Stalnaker-inspired presupposition theory ((Stalnaker, 1978)) and Clark-inspired grounding theory ((Clark, 1996)). Hence, KOS posits, following work in the tradition of dialogue games that conversational rules involve updates by each CP of her own *dialogue-gameboard* (DGB), a quasi-public informational repository (cf. Hamblin’s *individual commitment slate*, (Hamblin, 1970)). This allows conversational action to be viewed as operating on a publically accessible domain which is relative to each CP, and so parametrizable by unpublicized factors such as individual goals and intentions.

In KOS the DGB is construed as a data structure comprising the following attributes: **FACTS**: a set of facts corresponding to the information taken for granted by the CPs; **QUD** (‘questions under discussion’): a set consisting of the currently discussable questions, partially ordered

by \prec ('takes conversational precedence'); LATEST-MOVE: a linguistic sign whose content consists of the *latest move* made:

$$(13) \quad \left[\begin{array}{ll} \text{FACTS} & \text{set of facts} \\ \text{LATEST-MOVE} & \text{sign} \\ \text{QUD} & \text{p.o. set of questions} \end{array} \right]$$

Both querying and assertion involve a question becoming maximal in the querier/asserter's QUD: the posed question q for a query where q is posed, the polar question $p?$ for an assertion where p is asserted. Given this, we can define adjacency pair relations for dialogue moves: an adjacency pair to a query which poses q is a q -specific utterance; an adjacency pair to an assertion p is either a $p?$ -specific utterance or an acceptance move, in which a CP updates her FACTS structure with $\text{fact}(p)$.¹¹ A dialogue participant can downdate $q/p?$ from QUD when, as far as her (not necessarily public) goals dictate, sufficient information has been accumulated in FACTS. As an example, we can offer the following (simplified) analysis to the dialogue in (14):

- (14) A(1): Who's coming tomorrow?
 B(2): Several colleagues of mine.
 A(3): I see.
 B(4): Mike is coming too.

A's initial query in which she poses the question q_1 causes an update in her QUD. B takes up the question and, hence, updates his QUD. This gives him the opportunity to respond by asserting a proposition p_1 which provides information about q_1 . QUD is thus updated with the issue $p_1?$, which becomes the maximal element in QUD. A accepts the assertion, thereby updating her FACTS with the fact corresponding to p_1 and downdating $p_1?$ from QUD.¹² q_1 becomes maximal in QUD again, which licenses providing more information about this question, as B does in (4).¹³

¹¹If q is a question, a q -specific utterance is defined as follows: it is an utterance which either provides information σ About q or poses a question q_1 on which q Depends. Here About is partial answerhood and Depends is a relation between questions, which intuitively corresponds to the notion of 'is a subquestion of'. For more on these relations see (Ginzburg and Sag, 2000).

¹²This discussion is based on the assumption that facts and propositions are ontologically distinct. This is inessential for current purposes, in which case $\text{fact}(p_1)$ would simply be p_1 .

¹³We omit throughout here the specifications for LATEST-MOVE.

- (15) (1): $A \mid \text{QUD} := q_1$
 (2): $B \mid \text{QUD} : q_1$;
 asserts p_1 About q_1 ;
 $B \mid \text{QUD} := q_1 \prec p_1$?
 (3): $A \mid \text{QUD} := q_1 \prec p_1$?
 accepts p_1 ;
 $A \mid \text{FACTS} := \text{fact}(p_1)$;
 Downdates p_1 ? from $\text{QUD} := A \mid \text{QUD} : q_1$
 (4): given A 's acceptance:
 $B \mid \text{FACTS} := \text{fact}(p_1)$;
 downdates p_1 ? from $\text{QUD} : B \mid \text{QUD} := q_1$;
 asserts p_3 About q_1 ;
 $B \mid \text{QUD} := q_1 \prec p_3$?

3.2 TOPICALITY

In (Ginzburg, 1997b) I argued for the need to recognize a (TOPICAL v. STORED) dichotomy between two types of “presupposed informational items”: the former enter into FACTS via a “short-term”, defeasible repository structured by the elements of QUD. The latter are facts that either (after potentially being discussed) have been stored long-term or arise as side effects to issues that were never explicitly under discussion. This proposed dichotomy is inspired in part by work on fact and propositional anaphora in texts by (Webber, 1991; Asher, 1993) and enables one to propose a purely semantic, dialogical version of the Right Frontier Constraint (RFC) proposed for texts, based on an analogy that relates the text-derived notion of *open constituent* (“unexhausted topic”) with the dialogue-derived notion of *question currently under discussion*.

Formally, the distinction is captured by positing two distinct repositories within FACTS: STORED is modelled classically as a set of facts closed under meets and joins. TOPICAL is treated as a set of pairs of $a = \langle \text{question}_0, \text{fact}_0 \rangle$, where question_0 (a’s *address*) is an element of QUD, fact_0 is ABOUT q_0 . TOPICAL is updated using *priority union* (Carpenter, 1993), a defeasible update operation in which later accepted material takes precedence, hence allowing for an account of **hasty accommodation**. As far as querying and assertion: in this revised setup updating QUD has the additional consequence of introducing a new ADDRESS in TOPICAL about which facts can be provided, together initially with the trivial fact \top . In addition, when a new question gets introduced, the addresses for questions that are no longer under discussion are downdated from TOPICAL. This latter assumption represents a

simplified version of the RFC, which is generalized in (Ginzburg, 2001a). In particular, TOPICAL regulates the antecedents for bare *why*.

- (16) **It is precisely the facts in FACTS | TOPICAL to whom access by ellipsis and pronominal anaphora is possible.**

3.3 Utterance Representation: grounding and clarification

I start by offering an informal description of how an utterance u such as (17) can get grounded or spawn a clarification by an addressee B:

- (17) A: Did Bo leave?

A is attempting to convey to B her question whether the property she has referred to with her utterance of *leave* holds of the person she has referred to with the name *Bo*. B is required to try and find values for these references. Finding values is, with certain caveats, a necessary condition for B to ground A's utterance, thereby signalling that its content has been integrated in B's IS. Modelling this condition for successful grounding provides one obvious constraint on the representation of utterance types: such a representation must involve a function from or λ -abstract over a set of certain parameters (the *contextual parameters*) to contents. This much is familiar already from early work on context dependence by (Montague, 1974) et seq, who called such representations *meanings*. I will refer to the result of anchoring or instantiating a meaning of an utterance within an information state of a speaker as the *utterer's content*.

What happens when B cannot or is at least uncertain as to how he should instantiate in his IS a contextual parameter i ? In such a case B needs to do at least the following: (1) perform a partial update of the existing context with the successfully processed components of the utterance (2) pose a clarification question that involves reference to the sub-utterance u_i from which i emanates. Since the original speaker, A, can coherently integrate a clarification question once she hears it, it follows that, for a given utterance, there is a predictable range of <partial updates + consequent clarification questions>. These we take to be specified by a set of *coercion operations* on utterance representations.¹⁴ Indeed we assume that a component of dialogue competence is knowledge of these *coercion operations*.

¹⁴The term *coercion operation* is inspired by work on utterance representation within a type theoretic framework reported in (Cooper, 1998).

CE gives us some indication concerning both the input and required output of these operations. One such operation, **parameter identification**, essentially involves as output a question paraphrasable as *what is the intended reference of sub-utterance u_i ?* The partially updated context in which such a clarification takes place is such that simply repeating the segmental phonology of u_i using a distinctive intonation (e.g. focus-associated rise with spreading high tone or focus-associated fall with spreading low tone) enables that question to be expressed. Another existent coercion operation is **parameter focussing**. This involves a (partially updated) context in which the issue under discussion is a question that arises by instantiating all contextual parameters except for i and abstracting over i . In such a context, one can *confirm* that i gets the value B suspects it has by appropriately intoning any apparently co-referential phrase whose syntactic category is identical to u_1 's.

From this discussion, it becomes clear that coercion operations (and by extension the grounding process) cannot be defined simply on meanings. Rather, given the syntactic and phonological parallelism encoded in clarification contexts, these operations need to be defined on representations that encode in parallel for each sub-utterance down to the word level phonological, syntactic, semantic, and contextual information. With some minor modifications, signs as conceived in HPSG are exactly such a representational format and, hence, we will use them to define coercion operations.¹⁵ More precisely, given that an addressee might not be able to come up with a unique or a complete parse, due to lexical ignorance or a noisy environment, we need to utilize some 'underspecified' entity (see e.g. (Milward, 2000)). For simplicity we will use

¹⁵We make two minor modifications to the version of HPSG described in (Ginzburg and Sag, 2000)). First, we revamp the existing treatment of the feature C-INDICES. This will now encode the entire inventory of contextual parameters of an utterance (proper names, deictic pronouns, indexicals) not merely information about speaker/hearer/utterance-time, as standardly. Indeed, in principle, relation names should also be included, since they vary with context and are subject to clarification as well. Such a step involves a significant change to how argument roles are handled in existing HPSG. Hence, we do not make such a move here, though see (Purver, 2002) for a proposal. This modification of C-INDICES will allow signs to play a role akin to the role associated with 'meanings', i.e. to function as abstracts with roles that need to be instantiated. The second modification we make concerns the encoding of phrasal constituency. Standardly, the feature DTRS is used to encode immediate phrasal constituency. To facilitate statement of coercion operations, we need access to all phrasal constituents—given that a contextual parameter emanating from deeply embedding constituents are as clarifiable as immediate constituents. We posit a set valued feature CONSTIT(UE)NTS whose value is the set of all constituents immediate or otherwise of a given sign (Cf. the mother-daughter predicates used in (Gregory and Lappin, 1999).) In fact, having posited CONSTITS one could eliminate DTRS: this by making the value of CONSTITS be a set of sets whose first level elements are the immediate constituents. For current purposes, we stick with tradition and tolerate the redundancy of both DTRS and CONSTITS.

descriptions of signs. An example of the format for signs we employ is given in (18):¹⁶

$$(18) \left[\begin{array}{l} \textit{root-cl} \\ \text{PHON } \text{did bo leave} \\ \text{CAT } \text{V}[+\text{fin}] \\ \text{C-INDICES } \{ \boxed{1}, \boxed{2}, \boxed{3}, i, j \} \\ \text{CONT } \left[\begin{array}{l} \text{ASK-REL} \\ \text{ASKER } i \\ \text{ASKED } j \\ \text{MSG-ARG } \left[\begin{array}{l} \textit{question} \\ \text{PARAMS } \{ \} \\ \text{PROP} \mid \text{SOA } \left[\begin{array}{l} \textit{leave-rel} \\ \text{AGT } \boxed{1} \\ \text{TIME } \boxed{2} \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{CTXT} \mid \text{BCKGRD } \left\{ \textit{utt-time}(\boxed{3}), \right. \\ \left. \textit{precede}(\boxed{2}, \boxed{3}), \textit{named}(\text{bo})(\boxed{1}) \right\} \\ \text{CONSTITS } \left\{ \boxed{4}[\text{PHON Did}], \boxed{5}[\text{PHON Bo}], \right. \\ \left. \boxed{6}[\text{PHON leave}], \boxed{7}[\text{PHON Did Bo leave}] \right\} \end{array} \right]$$

3.4 Integrating Utterances in Information States

Before exemplifying one of the coercion operations, I need to explain how on our view utterances get integrated in an agent's IS. The basic protocol we assume is given in (19) below.¹⁷

(19) Utterance processing protocol

For an agent B with IS I : if an utterance u is Maximal in PENDING:

(a) Try to:

(1) find an assignment f in I for σ , where σ is the (maximal description available for) the sign associated with u

(2) update LATEST-MOVE with u :

¹⁶Within the phrasal type system of (Ginzburg and Sag, 2000) *root-cl* constitutes the 'start' symbol of the grammar. In particular, phrases of this type have as their content an illocutionary operator embedding the appropriate semantic object (an assertion embedding a proposition, a query embedding a question etc.). Here and throughout we omit various features (e.g. STORE, SLASH etc that have no bearing on current issues wherever possible).

¹⁷In this protocol, PENDING is a stack whose elements are (unintegrated) utterances.

- 1 If LATEST-MOVE is grounded, then $\text{FACTS} := \text{FACTS} + \text{LATEST-MOVE}$;
- 2 $\text{LATEST-MOVE} := \langle \sigma, f \rangle$
- (3) React to content(u) according to querying/assertion protocols.
- (4) If successful, u is removed from PENDING
- (b) Else: Repeat from stage (a) with MAX-QUD and SAL-UTT obtaining the various values of $\text{coe}^i(\tau) | \text{max} - \text{qud} / \text{sal} - \text{utt}$, where τ is the sign associated with LATEST-MOVE and coe^i is one of the available coercion operations;
- (c) Else: make an utterance appropriate for a context such that MAX-QUD and SAL-UTT get values according to the specification in $\text{coe}^i(u, \sigma)$, where coe^i is one of the available coercion operations.

The protocol involves the assumption that an agent always initially tries to integrate an utterance by assuming it constitutes an adjacency pair with the existing LATEST-MOVE. If this route is blocked somehow, because the current utterance cannot be grounded or the putative resolution leads to incoherence, only then does she try to repair by assuming the previous utterance is a clarification generated in accordance with the existing coercion operations. If that too fails, then, she herself generates a clarification. Thus, the prediction made by this protocol is that A will tend to initially interpret (20(2)) as a response to her question, not as a clarification:

- (20) A(1): Who do you think is the only person that admires Mary?
 B(2): Mary?

3.5 Sign Coercion

I now turn to exemplify one of the coercion operations we specified informally in section 3.3, namely **parameter identification**: for a given problematic contextual parameter its output is a partial specification for a sign whose content and MAX-QUD involve a question querying what the speaker intended to convey with the utterance requiring clarification:

(21) **parameter identification_i**:

$$\begin{aligned}
& \left[\begin{array}{l} \text{root-cl} \\ \text{C-PARAMS } \{ \dots \boxed{1} \dots \} \\ \text{CONSTITITS } \{ \dots \boxed{2} [\text{CONT } \boxed{3}] \dots \} \\ \dots \end{array} \right] \\
& \Rightarrow \\
& \left[\begin{array}{l} \text{CONTENT|MSG-ARG } \left[\begin{array}{l} \text{question} \\ \text{PROP } \boxed{3} \end{array} \right] \\ \text{C-PARAMS } \left\{ \dots \left[\begin{array}{l} \text{INDEX } \boxed{1} \\ \text{RESTR } \{ \text{addr}(\boxed{1}) \} \end{array} \right] \dots \right\} \\ \text{SAL-UTT } \boxed{2} \\ \text{MAX-QUD } \left[\begin{array}{l} \text{question} \\ \text{PARAMS } \{ [\text{INDEX } \boxed{4}] \} \\ \text{PROP } \boxed{3} \left[\begin{array}{l} \text{SOA} \left[\begin{array}{l} \text{utterer-cont-rel} \\ \text{AGENT } \boxed{1} \\ \text{SIGN } \boxed{2} \\ \text{CONT } \boxed{4} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]
\end{array} \right]
\end{aligned}$$

To exemplify: when this operation is applied to (18), it will yield as output the partial specification in (22):

$$(22) \left[\begin{array}{l} \text{CONT|MSG-ARG } \left[\begin{array}{l} \text{question} \\ \text{PROP } \boxed{6} \end{array} \right] \\ \text{C-PARAMS } \left\{ \dots \left[\begin{array}{l} \text{INDEX } \boxed{1} \\ \text{RESTR } \{ \text{addr}(\boxed{1}) \} \end{array} \right] \dots \right\} \\ \text{SAL-UTT } \boxed{5} \left[\begin{array}{l} \text{PHON bo} \\ \text{CAT NP} \\ \text{CONT|INDEX b} \end{array} \right] \\ \text{MAX-QUD } \left[\begin{array}{l} \text{question} \\ \text{PARAMS } \{ [\text{INDEX } \boxed{8}] \} \\ \text{PROP } \boxed{6} \left[\begin{array}{l} \text{SOA} \left[\begin{array}{l} \text{utterer-cont-rel} \\ \text{AGENT } \boxed{1} \\ \text{SIGN } \boxed{5} \\ \text{CONT } \boxed{8} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]
\end{array} \right]$$

(Ginzburg and Cooper,) provide detailed linguistic analysis showing how this specification allows for clarification questions such as the following:

- (23) a. Who do you mean Bo?
 b. WHO? (= who is Bo)
 c. Bo? (= who is Bo)

4. *Why_{meta}*: an analysis

4.1 Previous Work

There has been relatively little work on *Why_{meta}*, most of it by developers of dialogue systems. In MYCIN (Scott et al., 1984) *Why_{meta}* performed by the user is understood by the system to query *Why is the system trying to achieve* $\langle GOAL \rangle$, where $\langle GOAL \rangle$ is instantiated as the current higher level domain goal. (Moore, 1995) (see in particular section 5.2 pp. 180-199) provides a highly detailed and insightful analysis of the interpretation of bare *Why?*, which includes *Why_{meta}*. Her algorithm for interpreting a user's bare *Why?* works essentially as follows: the preferred interpretation is assumed to be *Why(p)*, where p is the most recent focus of attention (in the sense of (Sidner, 1981)) which is a constituent of the plan underlying the previous utterance. Other possible interpretations are less recent foci. The search for interpretations is bounded by the overall communicative goal that led to the system's utterance. Moore further postulates certain rules of inference associated with logical forms of the form *Why(illoc-prop)*, where *illoc-prop* ranges over logical forms of the form *Ask(A, q)*, *Inform(A, p)*, and *Order(A, a)*. These can be viewed as constituting the construals a bare *Why?* is assigned relative to the distinct types of utterance antecedents:¹⁸

- (24) a. *Why(Ask(system, q))*
 Why(Need – to – know(system, q))
 b. *Why(Order(system, do(user, a)))/*
 Why(do(user, a))
 c. *Why(Inform(system, user, p)))/*
 Why(believe(system, p))

Moore's proposal is clearly an insightful one and the resolutions it associates with *Why_{meta}* will give good results for a large number of

¹⁸Moore also notes, following observations by Van Fraassen and Garfinkel that bare *Why?* can also be focus sensitive, akin to adverbs like *only* and *even*. I will disregard this aspect of her account.

cases. A number of issues remain unresolved, nonetheless. For a start, no analysis is provided of the context which triggers *Why_{meta}*—the analysis provides the perspective of the system in trying to understand a user’s utterance of *Why_{meta}*, but not the perspective of someone wishing to *generate* uses of *Why_{metacommunicative}*. A related issue is providing a sufficiently general characterization of the triggering context to explain what other utterances are available. A third issue concerns adjacency: Moore’s account presupposes that *Why_{meta}* only occurs adjacently to the utterances whose clarification it seeks. As examples such as (5) indicate, this is not correct. This means we need an account as to which utterances constitute possible antecedents to a given use of *Why_{meta}* and how context is structured to enable such information to be accessed.

4.2 *Why_{meta}* as a metadiscursive utterance

I now turn to proposing my own analysis of *Why_{meta}*. One fundamental assumption I make, like Moore, is that *Why_{meta}* is simply a regular instance of a bare ‘Why’, which gets its particular content because the proposition/fact of which it predicates (concerning the content of an utterance) happens to be contextually available. Contextually available amounts to being *topical* in the sense explicated in section 3.2. In other words, it is a non-sentential utterance that satisfies the following schema:

$$(25) \left[\begin{array}{l} \text{CAT } \textit{adj} \\ \text{CONT } \left[\begin{array}{l} \textit{question} \\ \text{PROP } \left[\begin{array}{l} \textit{fact-mod(ifier)-rel} \\ \text{FACT-ARG } \boxed{1} \textit{ fact} \end{array} \right] \end{array} \right] \\ \text{CTXT|FACTS|TOPICAL } \left\{ \dots \boxed{1} \dots \right\} \end{array} \right]$$

The remaining task is to resolve the issues discussed in the final paragraph of the previous section: how is context structured to enable antecedents that give rise to *Why_{meta}* readings to emerge and disappear. The intuition I take as my starting point is that *Why_{meta}* is not a CR akin to those expressed by CE utterances in e.g. (1), (2) above, where information necessary for understanding the utterance is sought. Rather, *Why_{meta}* is a subspecies of utterances whose point is to consider whether a given issue warrants discussion—in short, metadiscursive utterances. Examples of these are given in (26):

(26) a. I don’t know.

b. Do we need to talk about this now?

c. I don't wish to discuss this now.

What we need to explain is how the potential for such reactions arises and how long it is maintained for. I assume that what is common to this class of reactions is that they more or less directly concern the issue of *whether B is willing to discuss q (p? for assertions)*¹⁹. Somewhat more pedantically: *whether at this point in the dialogue both CPs are interested for q (p? for assertions) to take conversational precedence*, often (but not always, cf. (26b)) entailing a negative resolution of this issue or raising an influencing issue. Let us call this issue 'QUD-Max(q)?'.

Thus, an important issue for a querier is whether her interlocuter intends to address the question posed or not.²⁰ And this is just about the first issue a *responder* is forced to consider once she has reasonable understanding of a query that has been addressed to her. In fact, one could envisage a protocol for querying in which posing *q* would involve an initial *discussion* of QUD-Max(q)?; if and once this is resolved positively, then *q* gets discussed. Such a view of querying, 'cautious querying', would be tantamount to thinking that all dialogues are implicitly as exemplified in (27):

(27) A: Can we now discuss this question, who left?

B: Yeah. Bill.

Similarly, one might take examples such as (28) to provide evidence for availability of QUD-Max(p?/q)? after both assertions and queries:²¹

(28) a. A: Ludwig was a lousy philosopher.

B: Yes/Right, no he wasn't. (cf. B: Yes! #no he wasn't.)

b. A: Was Ludwig a lousy philosopher?

B: Yes/Right, no he wasn't.

¹⁹I assume that *lack of ability to do X* precludes *intention to do X*.

²⁰I will concentrate here on *Why_{meta}* with query antecedents, for ease of exposition. However, the account is applicable to assertoric antecedents, as discussed below. Note that under certain conditions, the resolution of the discussion of *Why_{meta}* as a follow up to an assertion can lead to the original assertion getting discarded:

(i) A: A war is on the cards now.

B: Why (do you say that)?

A: There have been lots of planes in the air today.

B: Oh if that's your evidence I'm not even going to discuss the issue.

²¹These examples came up in a conversation with Staffan Larsson and Matt Purver.

In any case, **cautious querying** is not that far off the mark, as work in conversational analysis has suggested.²² However, we have to balance caution with another contextual fact: although by posing q A can not necessarily *impose* q as *maximal* in QUD, it still seems to be the case that q is added to his DGB. For a start we have exchanges like the following, where after an initial failure to accept, the querier seems to be able to use an elliptical short-answer form:

- (29) Cherrilyn: Are you still (pause) erm (pause) going to Bristol
(pause) on Monday?
Fiona: Dunno.
Cherrilyn: No? (BNC,)

We also want to be able to capture cases where the original querier keeps the turn to herself:

- (30) A: Anyway, so who is a good candidate? Maybe Bill, I guess.

The upshot of this is that we allow as an option for B to create a context where QUD-Max(q)? is the maximal element in QUD. Within such a context any QUD-Max(q)?-specific utterance may be uttered. Slightly more formally, we can posit a conversational rule whose output is the following partial specification for an utterance:

- (31)
$$\left[\begin{array}{ll} \text{MAX-QUD} & \text{QUD-Max}(q)? \\ \text{TOPICAL} & \{ \dots < \text{QUD-Max}(q)?, \text{Ask}(A, q) > \} \end{array} \right]$$

B introduces QUD-Max(q)? (= the issue (A and B interested-to-discuss q)?) into QUD; given that A has asked q , which we construe as A's interest to discuss q , this constitutes information about QUD-Max(q)? (i.e. is a partial answer to this question.), which is why TOPICAL gets the specification it does. Given this, *Ask*(A, q) is now available as an antecedent for fact anaphora/ellipsis, e.g. *Why_{meta}* or pronouns such as *this/that*.²³ A number of issues remain to be settled with respect to this rule, two above all: first, what is its triggering context; second, what is the status of q , the question A posed, in B's DGB. The initial thought one might have about the triggering context should be simply:

- (32) LATEST-MOVE: A Ask q

²²See, for instance, the CA notions of *pre-request* and *pre-announcement*, as discussed in (Levinson, 1983), p. 345 ff.

²³Examples

Such a specification flies in the face of examples such as (5) above. An obvious alternative, which allows for non-adjacency, as well as down-dating antecedents, would be to assume the triggering context is:

(33) MAX-QUD: q

This specification works reasonably well for examples such as the ones we've seen above. One apparent problem though is with reactions such as (26) above, where B does not respond at all to A's query, in which case one might doubt the motivation for assuming that q has made it into B's QUD. In other words, one might think that a consequence of a responder's failure to accept q for discussion is that q will only resurface if explicitly reposed. There is evidence, however, that actually q remains in a CP's QUD even when not initially adopted, its very posing makes it temporarily DGB available:

(34) A: Who are you meeting next week?
 B(2): I don't want to discuss this.
 A: Why?
 B: Personal reasons.
 A: I see.
 B(6): Oh, ok, Jill.

Here the original question has definitely *not* been reposed and yet B still has the option to address it, which he should be unable to do if it is not added to his gameboard before (34(2)). Thus, we can assume as a monotonic effect of querying that q enters the CP's QUD. In the case where metadiscussion occurs, QUD-Max(q)? gets pushed above q .

There are, however, a number of significant problems with choosing the triggering context to be QUD-maximality of q *if the elements in QUD are simply questions*. The first problem concerns a speaker/addressee asymmetry: as we noted previously, if A was the person posing a question q , then $why_{meta}(Ask(A, q))$ is simply unavailable to A at all subsequent points. This cannot be explained on 'pragmatic' grounds, since the speaker can fairly coherently express the requisite reading in non-elliptical fashion:

(35) a. Norrine: When is the barbecue, the twentieth? (pause) Something of June
 Chris: Thirtieth.
 Norrine: A Sunday.
 Chris: Sunday.
 Norrine: Mm. Why am I asking, eh?
 Chris: I don't know. Why?

The second problem concerns *why_{meta}* when the antecedent is an assertion, *Assert(A, p)*. In such a case *p?* does get into QUD. One might consider developing a strategy that manufactured resolutions of *why_{meta}* from the question in QUD along the following lines:

- (36) a. Antecedent is a wh-question *q*: *why_{meta}* is interpreted as *Why discuss q*
 b. Antecedent is a polar-question *p?*: *why_{meta}* is interpreted as *Why believe p*

Unfortunately, this will not always work, for the following reason: polar questions can enter QUD as a consequence of two types of dialogue moves, queries and assertions. But *why_{meta}* associated with a polar query does not always resolve identically to a *why_{meta}* associated with an assertion. (37a) shows one case where the paraphrase suggested in (36b) is somewhat plausible. However, (37b) shows, this is not invariably the case:

- (37) a. A: Are you ill? B: Why?
 b. [Context: Nancy has unexpectedly lost consciousness, a medic is attempting to resuscitate her:] A: Is Nancy going to be OK? B: Why?

It seems inescapable to conclude from these considerations that the antecedents for *why_{meta}* cannot merely be questions but have to be full contents of dialogue moves that include conversational move type (illocutionary force) information. In other words, propositions of the form in (38).²⁴ As we noted above, this is the content type clauses of type *root-cl* have in the framework of (Ginzburg and Sag, 2000).

²⁴It has been suggested to me by Jan van Kuppevelt that this is an overly restrictive assumption. He suggests that *why_{meta}* can also take a focus constituent of an utterance as an antecedent.

- (i) A: Will JOHN go on holidays this year?
 B: Why?/Why JOHN? (As you know, he has no money to go on vacation.)
 (ii) A: What's a good birthday present for Mary?
 B: A BIKE.
 A: Why?/Why a BIKE? (She got one two years ago and hardly used it)

It seems to me that in these examples the bare 'Why?' in these cases yields a regular *why_{meta}* interpretation, whose antecedent is the entire utterance (which might be a short answer). As for the cases where the bare 'why' takes an NP complement, these do not constitute *why_{meta}* interpretations. Thus, 'Why JOHN' does not have a reading paraphrasable as *Why are you asking if JOHN (of all people) is going on holiday..*

(38) a. *illoc-rel*(*spkr:A*, *msg-arg:message*)

b. *illoc-rel* ranges over *Ask*, *Assert*, *Order*, ... and *message* ranges over *question*, *proposition*, *outcome*, ...

There are a number of ways to implement this modification. Probably the most conservative would be to keep the architecture of the DGB as is and merely change the type of value that QUD takes. Instead of merely being questions, it would be pairs $\langle q, \text{illoc-prop} \rangle$ of questions and illocutionary propositions (the content of the move that gave rise to the question). The standard notion of QUD would involve the left projection of the enriched QUD value.^{25,26}

²⁵For those cases where QUD gets incremented indirectly, i.e. not as a consequence of an utterance, we can assume that the right projection is a trivial entity such as \perp .

²⁶A more radical move, proposed in an earlier draft of this paper, would be to have the value of QUD as simply the illocutionary proposition. Given a sufficiently structured notion of proposition, however, querying and assertion protocols more or less identical to the previously assumed ones are straightforward to formulate. One problem such a proposal faces is in capturing the contextual commonality exhibited by an assertion p and a query $p?$, which allows both to serve as antecedents of *yes*. In the existing account of QUD this is captured by assuming that both contribute $p?$ to QUD:

- (i) A: Bo is coming here tomorrow./ Is Bo coming here tomorrow? B: Yes/Indeed/Probably/No/No way
- (ii) A: Can I have one of those um/I'm going to have one of those um B: What things? A: The pills. B: Penicillin? A: Yes. B: They are a bit addictive. A: I realize. B: Yes, OK.

Thus, in (Ginzburg and Sag, 2000), the semantic/contextual part of the lexical entry for *yes* is as in (iii) below—the content is identified with the PROP ('open proposition' component) of the polar question that MAX-QUD is specified to be:

$$(iii) \left[\begin{array}{c} \text{CONT} \quad \boxed{1} \\ \text{CTXT} | \text{MAX-QUD} \quad \left[\begin{array}{cc} \text{PARAMS} & \{ \} \\ \text{PROP} & \boxed{1} \end{array} \right] \end{array} \right]$$

Independent motivation for the 'radical approach' would be minimal pairs, where the existing QUD representation was identical, but for which there is evidence that long-term (i.e. beyond adjacency, where moves are determined by LATEST-MOVE), the states need to be distinguished. Matt Purver has pointed out to me examples such as (iv,v), which might indeed suggest the need to distinguish the long-term contextual impact of an assertion p from a query $p?$:

- (iv) A(1): Is Bill coming to the party?
B(2): Is he back from the USA yet?
A(3): Yes, he got back last week.
B(4): then yes./ # oh, ok.
- (v) A(1): Bill is coming to the party.
B(2): Is he back from the USA yet?
A(3): Yes, he got back last week.
B(4): oh, ok./ #then yes.

An analysis of these cases remains elusive.

4.3 Formulating Metadiscursive Triggering

Assuming this modification done, both problems noted previously for a QUD-based triggering of metadiscussion are solved immediately: if the antecedent of why_{meta} includes also information about the identity of the speaker, then application of the conversational rule for generating metadiscursive responses can be restricted to the original responder to that utterance. Similarly, there is no difficulty distinguishing between antecedents of assertions and of queries whose descriptive content is a polar question.

We can now provide a fuller formulation of our rule for metadiscursive utterances: in a context where q_0 is QUD-maximal, where q_0 originates from an utterance whose speaker was A, B may alter his context to one in which QUD-Max(q_0)? is QUD-maximal. This makes the illocutionary fact concerning q_0 TOPICAL. B now provides a QUD-Max(q_0)?-specific utterance.

(39) **metadiscursive utterance rule:**

A. Input context:

$$\left[\begin{array}{l} \text{FACTS|STORED } \{ \text{spkr}(\text{B}), \text{addr}(\text{A}) \} \\ \text{QUD } \dots \prec \langle q_0, \text{Illoc-rel}(\text{A}, \text{r}) \rangle \end{array} \right]$$

B. Modified context:

$$\left[\begin{array}{l} \text{FACTS|TOPICAL } \{ \dots < \text{QUD-MAX}(q_0)?, \text{Illoc-rel}(\text{A}, \text{r}) > \} \\ \text{QUD } \dots \prec \langle q_0, \text{Illoc-rel}(\text{A}, \text{r}) \rangle \prec \langle \perp, \text{QUD-Max}(q_0)? \rangle; \end{array} \right]$$

C. B makes a QUD-Max(q_0)?-specific utterance

Let me exemplify this by considering two dialogues.

- (40) A(1): Who's coming tomorrow?
 B(2): Why?
 A(3): I need to plan.
 B(4): Oh, OK. Jill Wenderfleisch.

(41)

(1): A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$
 (2): B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle \perp, \text{QUD-Max}(q_1)? \rangle$;
 B|FACTS | TOPICAL: $\{ \langle \text{QUD-Max}(q_1)?, \text{Ask}(A, q_1) \rangle \}$;
 asks Why(Ask(A, q_1)) which influences QUD-Max(q_1)?;
 B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle \perp, \text{QUD-Max}(q_1)? \rangle$
 $\prec \langle \text{Why}(\text{Ask}(A, q_1)), \text{Ask}(B, \text{Why}(\text{Ask}(A, q_1))) \rangle$
 (3): A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle \perp, \text{QUD-Max}(q_1)? \rangle$
 $\prec \langle \text{Why}(\text{Ask}(A, q_1)), \text{Ask}(B, \text{Why}(\text{Ask}(A, q_1))) \rangle$
 asserts p_1 about Why(Ask(A, q_1));
 A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle \perp, \text{QUD-Max}(q_1)? \rangle$
 $\prec \langle \text{Why}(\text{Ask}(A, q_1)), \text{Ask}(B, \text{Why}(\text{Ask}(A, q_1))) \rangle \prec \langle p_1?, \text{Assert}(A, p_1) \rangle$;
 (4): accepts p_1 : B|FACTS | TOPICAL: $\{ \langle \text{QUD-Max}(q_1)?, \text{Ask}(A, q_1) \rangle, \langle p_1?, \text{fact}(p_1) \rangle, \langle \text{Why}(\text{Ask}(A, q_1)), \text{fact}(p_1) \rangle \}$;
 downdates $\langle \perp, \text{QUD-Max}(q_1)? \rangle, \langle \text{Why}(\text{Ask}(A, q_1)), \text{Ask}(B, \text{Why}(\text{Ask}(A, q_1))) \rangle, \langle p_1?, \text{Assert}(A, p_1) \rangle$ from QUD:
 B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$;
 asserts p_3 about q_1 ;

- (42) A(1): Who's coming tomorrow?
 B(2): Jill Wenderfleisch.
 A(3): I see.
 B(4): Mmh. Why? (= *Why did you ask who is coming or Why is Jill Wenderfleisch coming tomorrow*)²⁷

(43)

(1): A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$
 (2): B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$; asserts p_3 about q_1 :
 B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle p_3?, \text{Assert}(B, p_3) \rangle$;
 (3): A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle p_3?, \text{Assert}(B, p_3) \rangle$;
 accepts p_3 : A|FACTS | TOPICAL: $\{ \langle p_3?, \text{fact}(p_3) \rangle, \langle q_1, \text{fact}(p_3) \rangle \}$
 A | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$
 (4): accepts p_3 :
 B|FACTS | TOPICAL: $\{ \langle p_3?, \text{fact}(p_3) \rangle, \langle q_1, \text{fact}(p_3) \rangle \}$;
 downdates $\langle p_3?, \text{Assert}(B, p_3) \rangle$ from QUD:
 B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle$
To explain Why_{meta} reading the following update occurs
 B | QUD: $\langle q_1, \text{Ask}(A, q_1) \rangle \prec \langle \perp, \text{QUD-Max}(q_1)? \rangle$
 B|FACTS | TOPICAL: $\{ \langle q_1, \text{fact}(p_1) \rangle, \langle \text{QUD-Max}(q_1)?, \text{Ask}(A, q_1) \rangle \}$;
 asks Why(Ask(A, q₁)) which influences QUD-Max(q₁)?;

5. Concluding Remarks

The starting point of this paper was my claim that analyses of interaction need to characterize not solely ‘success conditions’, a traditional and important means of analyzing action, but also ‘clarification potential’, the range of potential clarification requests available in the aftermath of a conversational move. Indeed, as I exemplified, there are very productive and effective ways of producing CRs relating to the grammatically governed content of an utterance. Of course not all the content associated with a conversational move is grammatically based. Some is indirectly conveyed on the basis of a CP’s non-public intentions. Interestingly there exist also highly productive means of making CRs about this aspect of an interaction, one of the commonest means being by means of the bare *Why?*, the relevant reading of which I dubbed *Why_{meta}*. Nonetheless, I demonstrated that *Why_{meta}* shows distinct behaviour from CRs that pertain to grammatically governed content. The most prominent feature perhaps being that, whereas the latter are almost invariably adjacent to

²⁷The latter reading is forced if *why* is embedded, e.g. *I wonder why*.

the utterances whose clarification they seek, non-adjacency is quite natural for *Why_{meta}*. It can occur at a stage where a second part adjacency pair response has been provided to the utterance it pertains to, suggesting that the information *Why_{meta}* is seeking is a ‘useful extra’, not an essential ingredient required for providing an appropriate response.

Given this, I argued that the contextual means needed to trigger *Why_{meta}* should involve a distinct mechanism from that used to trigger CRs that concern grammatically governed content. Rather than treat *Why_{meta}* as clarifying a contextually instantiable goals/plan parameter, I proposed it be treated as an instance of a metadiscursive utterance like *I don’t want to talk about this*. The main features of the account of *Why_{meta}* I have offered are these:

- 1 *Why_{meta}* is treated with the same tools as other uses of *Why?*.
- 2 *Why_{meta}* is only available to the original addressee of the utterance.
- 3 *Why_{meta}* is available as long as the utterance it concerns is under discussion.

One thing that I have *not* done in this paper is provide a mechanism for explicating interaction over non-grammatically governed content. Within the current perspective this is most naturally seen as arising from reasoning about the unpublicized part of the other CP’s information state. An explicit and computationally implemented theory of the unpublicized part of an information state that complements the DGBs assumed here has been worked out by Staffan Larsson, Robin Cooper and the other developers of the GODIS system (Larsson, 2002; Cooper et al., 2000). The basic approach which seems most consistent with the observations provided in this paper is that communicative intentions are, to a large extent, in tight correspondence with information represented on the DGB. In other words, CPs do not on the whole keep track explicitly of the unpublicized part of the other CP’s information state, given an assumption of transparency between this and the DGB. For psycho-linguistic evidence supporting this conclusion see (Garrod and Pickering,). This is also supported by corpus results concerning the resolution of non-sentential utterances (Fernández and Ginzburg,), which suggest that non-linguistically provided material is used very sparingly in ellipsis resolution.

This is not to suggest that reasoning about the unpublicized part of the other CP’s information state does not or cannot take place, particularly when a CP is being circumspect, wishing to optimize cooperativity or politeness. Thus, in (44), taken from a conversation in a run down British university, Bo does not address the question posed by Anja, rea-

soning that she is primarily interested to leave the building, not learn about its geography. Given that his response is based on reasoning about why she asked the question, *Why_{meta}* is inappropriate here:

- (44) a. Anja: Where is the lift?
 Bo: We use the stairs mostly; the lift rarely works. #Why?

References

- Allen, J. and Perrault, R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178.
- Asher, N. (1993). *Reference to Abstract Objects in English: a Philosophical Semantics for Natural Language Metaphysics*. Studies in Linguistics and Philosophy. Kluwer, Dordrecht.
- Bohlin, P., Cooper, R., Engdahl, E., and Larsson, S. (1999). Information states and dialogue move engines. *Gothenburg Papers in Computational Linguistics*.
- Carpenter, B. (1993). Skeptical and credulous default unification with applications to templates and inheritance. In T. Briscoe, A. Copestake, V. d. P., editor, *Inheritance, Defaults, and the Lexicon*. Cambridge University Press, Cambridge.
- Clark, H. (1996). *Using Language*. Cambridge University Press, Cambridge.
- Cooper, R. (1998). Mixing situation theory and type theory to formalize information states in dialogue exchanges. In Hulstijn, J. and Nijholt, A., editors, *Proceedings of TwenDial 98, 13th Twente workshop on Language Technology*. Twente University, Twente.
- Cooper, R., Larsson, S., Hieronymus, J., Ericsson, S., Engdahl, E., and Ljunglof, P. (2000). *GODIS and Questions Under Discussion*. University of Gothenburg, Gothenburg. Available from <http://www.ling.gu.se/research/projects/trindi>.
- Fernández, R. and Ginzburg, J. Non-sentential utterances: a corpus-based study. *Traitement Automatique des Langues*. Forthcoming.
- Garrod, S. and Pickering, M. Toward a mechanistic psychology of dialogue. *Behavioural and Brain Sciences*. Forthcoming.
- Ginzburg, J. (1996). Interrogatives: Questions, facts, and dialogue. In Lappin, S., editor, *Handbook of Contemporary Semantic Theory*. Blackwell, Oxford.
- Ginzburg, J. (1997a). On some semantic consequences of turn taking. In Dekker, P., Stokhof, M., and Venema, Y., editors, *Proceedings of the 11th Amsterdam Colloquium on Formal Semantics and Logic*, pages 145–150. ILLC, Amsterdam.

- Ginzburg, J. (1997b). Structural mismatch in dialogue. In Jaeger, G. and Benz, A., editors, *Proceedings of Mundial 97 (Technical Report 97-106)*, pages 59–80. Universitaet Muenchen Centrum fuer Informations- und Sprachverarbeitung, Muenchen.
- Ginzburg, J. (2001a). Clarification ellipsis and nominal anaphora. In Bunt, H., Muskens, R., and Thijsse, E., editors, *Computing Meaning: Volume 2*, number 77 in Studies in Linguistics and Philosophy. Kluwer, Dordrecht.
- Ginzburg, J. (2001b). Turn taking puzzles and the semantics of adjuncts. Paper presented at Informatics Faculty Seminar, October 26, 2001.
- Ginzburg, J. (forthcoming). *Semantics and Interaction in Dialogue*. CSLI Publications and Cambridge University Press, Stanford: California. Draft chapters available from <http://www.dcs.kcl.ac.uk/staff/ginzburg>.
- Ginzburg, J. and Cooper, R. Clarification, ellipsis, and the nature of contextual updates. *Linguistics and Philosophy*. to appear.
- Ginzburg, J. and Cooper, R. (2001). Resolving ellipsis in clarification. In *Proceedings of the 39th Meeting of the Association for Computational Linguistics*.
- Ginzburg, J. and Sag, I. A. (2000). *Interrogative Investigations: the form, meaning and use of English Interrogatives*. Number 123 in CSLI Lecture Notes. CSLI Publications, Stanford: California.
- Gregory, H. and Lappin, S. (1999). Antecedent contained ellipsis in hpsg. In Webelhuth, G., Koenig, J. P., and Kathol, A., editors, *Lexical and Constructional Aspects of Linguistic Explanation*, pages 331–356. CSLI Publications, Stanford.
- Hamblin, C. L. (1970). *Fallacies*. Methuen, London.
- Larsson, S. (2002). *Issue based Dialogue Management*. PhD thesis, Gothenburg University.
- Levinson, S. (1983). *Pragmatics*. Cambridge University Press, Cambridge.
- Milward, D. (2000). Distributing representation for robust interpretation of dialogue utterances. *ACL*.
- Montague, R. (1974). Pragmatics. In Thomason, R., editor, *Formal Philosophy*. Yale UP, New Haven.
- Moore, J. (1995). *Participating in Explanatory Dialogues*. Bradford Books. MIT Press, Cambridge.
- Purver, M. (2001). Score: A tool for searching the bnc. Technical report, King’s College, London.
- Purver, M. (2002). Processing of unknown words in a dialogue system. In *SigDial Workshop on Discourse and Dialogue*.
- Purver, M., Ginzburg, J., and Healey, P. On the means for clarification in dialogue. In *(This Volume)*.

- Scott, C., Clancey, W., Davis, R., and Shortliffe, E. (1984). Methods for generating explanations. In Buchanan, B. and Shortliffe, E., editors, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, pages 338–362. Addison Wesley.
- Sidner, C. (1981). Focusing for interpretation of pronouns. *American Journal of Computational Linguistics*, 4:217–231.
- Stalnaker, R. C. (1978). Assertion. In Cole, P., editor, *Syntax and Semantics, Volume 9*, pages 315–332. AP, New York.
- Webber, B. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, 14:107–135.